DØ Note 5353-CONF

A Search for the Standard Model Higgs Boson in the Channel $ZH \to \nu\overline{\nu}b\overline{b}$ at $\sqrt{s} =$ 1.96 TeV

The DØ Collaboration
URL http://www-d0.fnal.gov
(Dated: April 12, 2007)

This note describes a search at DØ for the Standard Model Higgs boson produced in association with a Z boson, where the Higgs decays to a $b\overline{b}$ pair and the Z boson decays to neutrinos. An integrated luminosity of 0.930 fb^{-1} was used. The $p\overline{p} \to ZH \to \nu\overline{\nu}b\overline{b}$ channel is one of the most sensitive ways to search for a light mass Higgs, but is problematic at hadron colliders due to the absence of visible leptons and the presence of only two jets in the final state. We thus require a large missing transverse energy and that the jets be b-tagged, with any Higgs signal producing a peak in their invariant mass. This analysis updates the previously published DØ analysis in this channel [1]. In the absence of a significant excess we set limits on $\sigma(p\overline{p} \to ZH) \times (H \to b\overline{b})$ at the 95% confidence level of 2.7 pb – 1.6 pb for Higgs boson masses from 105 – 135 GeV.

Preliminary Results for Winter 2007 Conferences

I. INTRODUCTION

The Higgs mechanism is the best candidate to explain electroweak symmetry breaking. It predicts the existence of the Higgs boson, which is as yet undiscovered. The LEP experiments have placed a lower limit on its mass at 114.4 GeV at 95% confidence level (CL) [2]. Electroweak global fits prefer a light mass Higgs, currently $m_H < 144$ GeV at 95% CL [3]. At these masses, the Tevatron has significant discovery potential, details of which can be found elsewhere [4][5].

The $p\overline{p} \to ZH \to \nu\overline{\nu}b\overline{b}$ channel is one of the most sensitive ways to search for a light mass Higgs because of the large $H \to b\overline{b}$ and $Z \to \nu\overline{\nu}$ branching ratios. However the absence of visible leptons and the presence of only two jets in the final state mean that we must require a large missing transverse energy ($\not\!\!E_T$) and that the jets be b-tagged. The two b-jets from the Higgs are boosted along the direction of the Higgs momentum and so tend to be more acoplanar than the dijet background. There are two major sources of background: (i) physics backgrounds such as Z+jets, W+jets, electroweak diboson production or top quark production with missed leptons and jets and (ii) the instrumental background resulting from calorimeter mismeasurements which can lead to high $\not\!\!E_T$ signals with the presence of jets from QCD processes. Selecting events with a relatively large $\not\!\!E_T$ and two b-tagged jets eliminates much of the physics background.

This analysis updates DØ's previously published result in this channel [1]. As well as more data, this analysis benefits from improved b-tagging using a neural net, and improved calorimeter calibration.

II. DATA SAMPLE AND EVENT SELECTION

The Run II DØ detector has a central-tracking system, consisting of a silicon microstrip tracker and a central fiber tracker, both located within a 2 T superconducting solenoidal magnet, with designs optimised for tracking and vertexing at pseudorapidities $|\eta| < 3$ and $|\eta| < 2.5$, respectively. The liquid-argon and uranium calorimeter has a central section (CC) covering pseudorapidities $|\eta|$ up to ≈ 1.1, and two end calorimeters (EC) that extend coverage to $|\eta| \approx 4.2$, with all three housed in separate cryostats. An outer muon system, at $|\eta| < 2$, consists of a layer of tracking detectors and scintillation trigger counters in front of 1.8 T toroids, followed by two similar layers after the toroids. Luminosity is measured using plastic scintillator arrays placed in front of the EC cryostats. Full details of the Run II DØ detector are given elsewhere [6].

Dedicated triggers designed to select events with acoplanar jets and large missing E_T were used. After data quality cuts the total data sample is 0.930 fb^{-1} [7]. As the data were taken with different trigger versions, the exact criteria vary, but typical requirements at the highest level trigger (Level-3) were a vectorial $\not\!\!H_T > 30$ GeV (where H_T is the sum of the jet p_T) and an azimuthal angle, ϕ, between the two leading P_T jets of $\phi < 170°$. A parameterised trigger simulation was used to model the effects of the trigger requirements on the simulated events. Each trigger version was simulated separately and the different versions combined with the appropriate luminosity weighting.

The basic event selection is as follows:

- At least two jets with $P_T > 20$ GeV and $|\eta| < 1.1$ or $1.4 < |\eta| < 2.5$ [13].
- $\Delta\phi(jet_1, jet_2) < 165°$.
- $\not\!\!E_T > 50$ GeV.
- The presence of a primary vertex with $|z| < 35$ cm, with at least 3 attached tracks.
- $H_T < 240$ GeV.
- No isolated leptons (electron or muon) [14].

The first three requirements select the basic signal topology and requiring a good vertex maximises the b-tagging discrimination potential. The last two requirements are designed to reduce the contribution of $t\overline{t}$ events; if the W from the top quark decays to jets then the total scalar sum of the jets in the events will be large and conversely if the W decays leptonically then the isolated lepton cut will reject these. The isolated lepton cut also eliminates a significant proportion of the analysed leptonic decays of Z (non-neutrino) and W bosons. The jets are required to pass basic quality cuts to remove fake jets. Corrections are applied as a function of η and p_T to correct the jet energies back to the particle level for detector and physics effects. As in Ref [1] additional requirements are made to reduce the instrumental background and are detailed in Section IV.

III. SIMULATED EVENT SAMPLES

The samples listed below were used to determine the number of expected signal and background events:

- Signal samples, $ZH \to \nu\nu b\bar{b}$, $WH \to e\nu_e b\bar{b}$, $WH \to \mu\nu_\mu b\bar{b}$ and $WH \to \tau\nu_\tau b\bar{b}$ were generated for Higgs masses, m_H, from 105 to 135 GeV using PYTHIA 6.323 [8]. The latter signals are included because if the charged lepton from the W decay escapes then the WH production process will contribute to the signal in this $\not\!\!E_T$ + jets topology and so is considered part of the signal.
- $t\bar{t}$ pair production with up to 4 jets, generated with ALPGEN v2.05 [9].
- W+jets (including jj, $b\bar{b}$ and $c\bar{c}$ jets separately) and Z+jets (including $Z \to \nu\overline{\nu}$ and $Z \to \tau\tau$ processes for jj, $b\bar{b}$ and $c\bar{c}$ jets) samples were generated using ALPGEN.
- Diboson processes, namely WW, WZ and ZZ, were generated with PYTHIA.

Those samples generated with ALPGEN were passed through PYTHIA for showering and hadronisation. NLO cross sections were used for normalisation for all processes other than diboson production. All samples were passed through the DØ detector simulation, the readout simulation and the reconstruction software. Prior to b-tagging the jets are required to be 'taggable' i.e. to satisfy certain tracking and vertexing criteria. The fraction of taggable jets was investigated as a function of p_T, η and the z-position of the primary vertex (PV_Z) using a W+jets data sample [15] and Monte Carlo (MC). The scale factor, the ratio of the taggability in data and MC, was found to only depend on η. Scale factors of 0.97 ± 0.01 and 0.95 ± 0.03 where the uncertainty is statistical are used for the central and endcap regions respectively. Residual differences between data and simulation are taken into account by smearing the simulation.

IV. INSTRUMENTAL BACKGROUND ESTIMATION

The following additional requirements are introduced to reduce the instrumental background:

- The minimum of the difference in azimuthal angle, ϕ, between the direction of the $\not\!\!E_T$ and any of the jets, $min\Delta\phi(\not\!\!E_T, jets) > 0.15$, and the related cut, $\not\!\!E_T$ (GeV) $> -40 \times min\Delta\phi(\not\!\!E_T, jets) + 80$.
- The difference in azimuthal angle between the direction of $\not\!\!E_T$ and $p_T^{neg.\,trk.}$: $\Delta\phi(\not\!\!E_T, p_T^{neg.\,trk.}) < \pi/2$. $p_T^{neg.\,trk.}$ is the magnitude of the negative of the vector sum of all track p_Ts.
- The asymmetry between $\not\!\!E_T$ and $\not\!\!H_T$: $-0.1 < Asym(\not\!\!E_T, \not\!\!H_T) < 0.2$, where the $Asym(\not\!\!E_T, \not\!\!H_T) \equiv (\not\!\!E_T - \not\!\!H_T)/(\not\!\!E_T + \not\!\!H_T)$.

For events originating from hard scattering with genuine missing transverse energy the $\not\!\!H_T$, $\not\!\!E_T$ and $p_T^{neg.\,trk.}$ all point in the same direction and are correlated. By contrast, dijet events in which one of the jets has been mismeasured to give an $\not\!\!E_T$ signal have $\phi(\not\!\!E_T)$ in the same direction as one of the jets. The variables mentioned above were studied in W+jets data, an orthogonal sample to that used in the final analysis, to ensure that they are understood. Distributions of the p_T for the leading and next-to-leading jet, along with $\not\!\!E_T$ and $\Delta\phi(\not\!\!E_T, p_T^{neg.\,trk.})$ are shown in Fig. 1.

Whilst the requirements above reduce the instrumental background, its absolute scale still needs to be determined from data to calculate any residual contribution. The $Asym(\not\!\!E_T, \not\!\!H_T)$ and $\Delta\phi(\not\!\!E_T, p_T^{neg.\,trk.})$ distributions are used to make this estimate. The simulated signal and physics backgrounds peak around 0 for both variables. The signal region is defined as having $\Delta\phi(\not\!\!E_T, p_T^{neg.\,trk.}) < \pi/2$ and the sideband region as having $\Delta\phi(\not\!\!E_T, p_T^{neg.\,trk.}) > \pi/2$. To estimate the scale of the instrumental background we consider the $Aysm(\not\!\!E_T, \not\!\!H_T)$ distribution for the signal and sideband regions. The shape of the physics backgrounds is, for either region, taken directly from simulation. We fit a sixth-order order polynomial to the the $Aysm(\not\!\!E_T, \not\!\!H_T)$ distribution in the sideband region to determine the shape for the instrumental background (having subtracted the physics background contamination.) Thus, having then determined the shape of the instrumental $Aysm(\not\!\!E_T, \not\!\!H_T)$ distribution we then make a combined (Monte Carlo and instrumental) fit to data in the signal region, as shown in Fig. 2. The sixth-order polynomial fits the data well. For this combined fit the simulation and instrumental shapes are fixed to those from the previous fits and only the absolute scale of the two types of background is allowed to float. The normalisation of the physics background is found to be 1.08 ± 0.02 (statistical error only) in good agreement with the expected cross sections. For all the other distributions the instrumental background also has to have its shape adjusted for the physics background contamination in the sideband region. The data-simulation comparison for the p_T for the leading and next-to-leading jet, $\not\!\!E_T$ and dijet mass are shown after selection cuts, taggability corrections and the normalisation of the instrumental and physics backgrounds in Fig. 3.

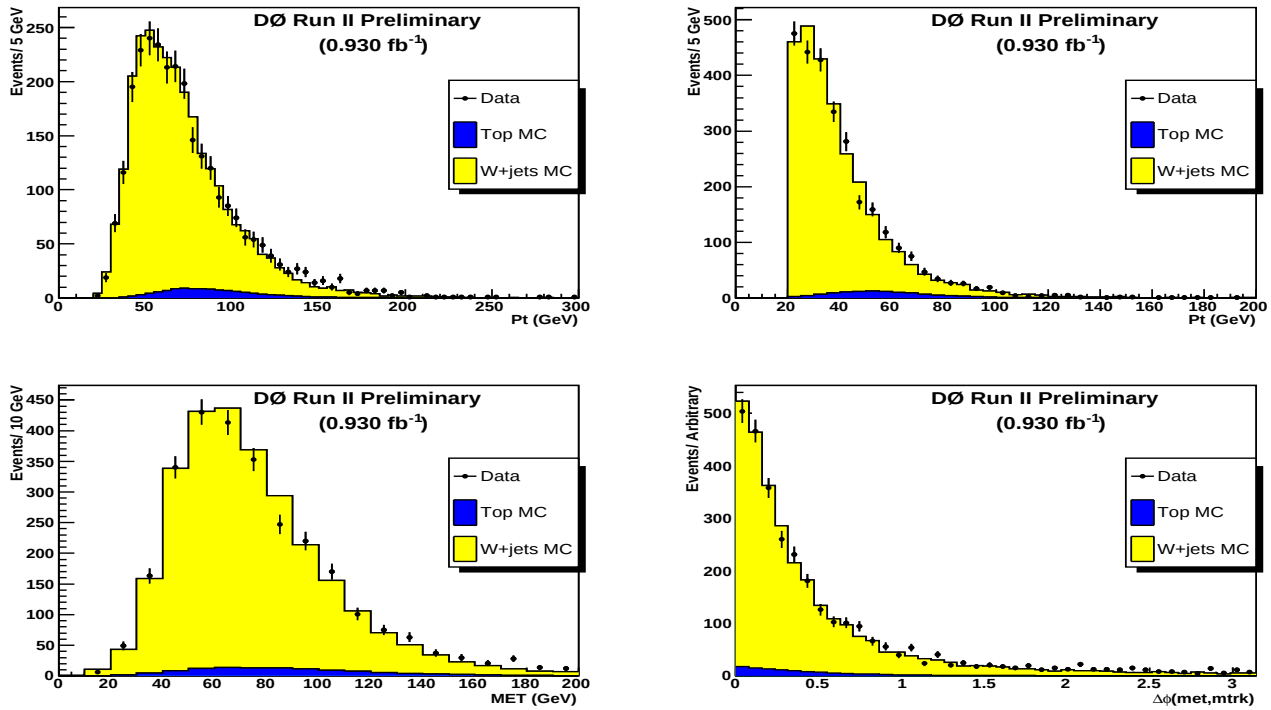

FIG. 1: Distributions of the p_T for the leading (upper left) and next-to-leading jet (upper right), $\not\!\!E_T$ and $\Delta\phi(\not\!\!E_T, p_T^{neg.\,trk.})$ for W+jets events before b-tagging.

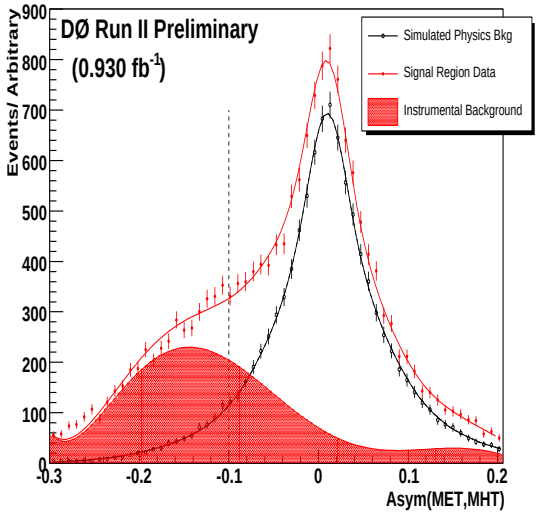

FIG. 2: $Asym(\not\!\!E_T, \not\!\!H_T)$ for data, physics background and instrumental background in the signal region before b-tagging. The selection cut is $-0.1 < Asym(\not\!\!E_T, \not\!\!H_T) < 0.2$, the line at -0.1 indicates the lower edge of this range.

V. BTAGGED RESULTS

The standard DØ neural net b-tagging algorithm was used [10]; this uses as input the output from several lifetime based b-tagging algorithms. As the (sideband) data are used to estimate the QCD / instrumental background the b-tagging is applied directly to the sideband data which naturally have limited statistics; this statistical error generates the systematic error referred to as 'Sideband b-tagging' in table II. The final selection uses the following optimum

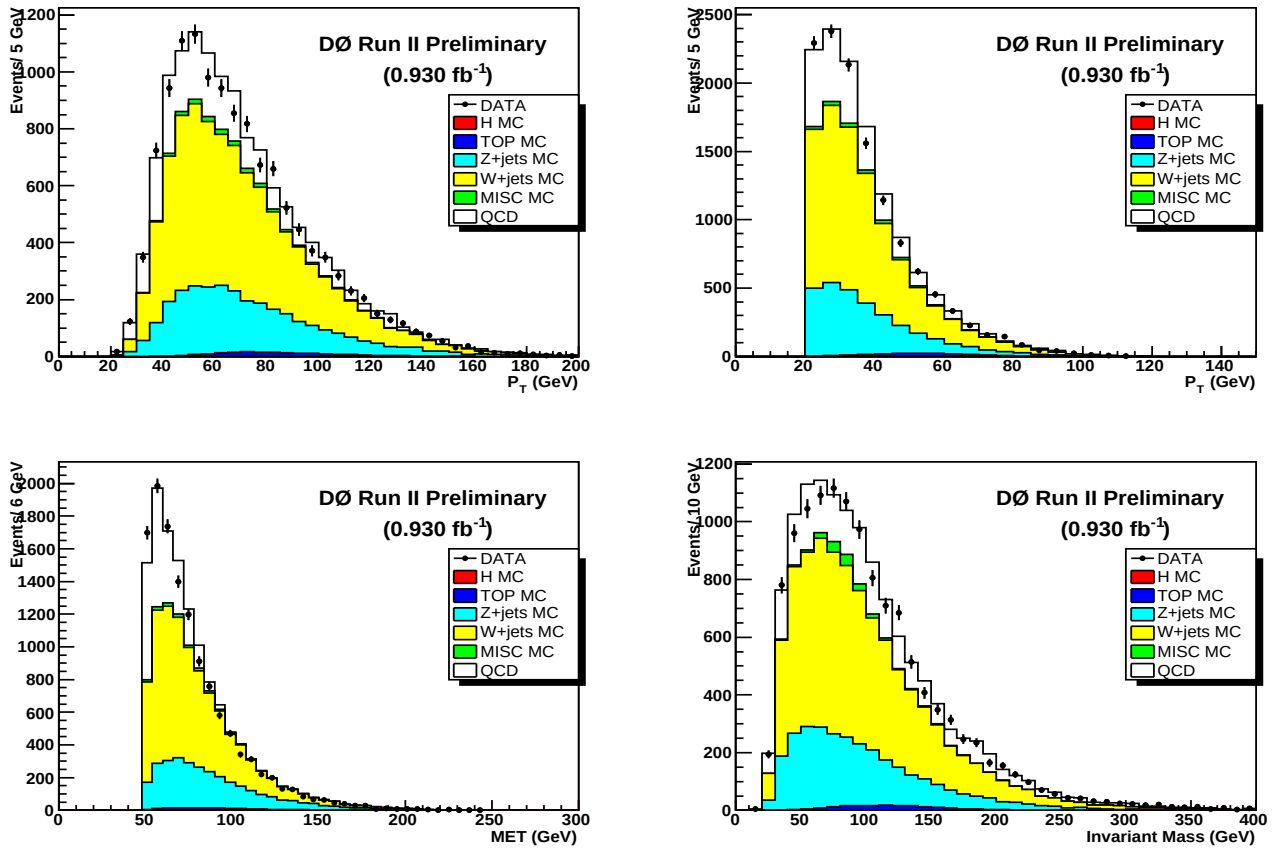

FIG. 3: Distributions of the p_T for the leading (upper left) and next-to-leading jet (upper right), $\not\!\!E_T$ and dijet mass before b-tagging.

b-tagging combination: one tight b-tag (b-tag efficiency ∼50% for a mistag rate of ∼0.4%) and one loose b-tag (b-tag efficiency ∼70% for a mistag rate of ∼4.5%). Table I shows the number of expected events from simulation and instrumental background, along with the number seen in data, before and after b-tagging requirements. After b-tagging 130.1 events are expected, and 140 events are observed. Fig. 4 shows the p_T for the leading and next-to-leading jet, $\not\!\!E_T$ and dijet mass after applying the optimised b-tagging.

VI. SYSTEMATIC ERRORS

Systematic errors associated with the trigger efficiencies, jet energy scale, normalisation procedure, QCD side band fitting procedure and b-tagging (including the QCD sideband) are estimated and included in the limit setting procedure. The overall signal and background errors are 15% and 14% respectively. Errors were evaluated by varying each source of uncertainty by $\pm 1\sigma$ and repeating the analysis. The normalisation error is the combination in quadrature of the error in the normalisation fit and the error on the individual cross section under question. Table II lists the systematic errors.

VII. CROSS SECTION LIMITS

We apply a mass window on the double b-tagged dijet invariant mass of $\pm 2\sigma$ around the mean, where σ is the estimated Gaussian width of the $H \to b\bar{b}$ invariant mass peak, as optimised from the simulation. The mass resolution in the region of interest is around 17%. The numbers of events expected and observed for signal and background are given in Table III. For the 115 GeV Higgs mass window, we observe 75 events for an estimated background of 63.3

Sample	No b-tag	Double b-tag
$ZH(m_H = 115\ GeV)$	2.52	0.90
$WH(m_H = 115\ GeV)$	1.79	0.63
$W + jets$	6905	53.5
$\rightarrow Wjj$	5302	7.73
$\rightarrow Wbb$	406.1	36.2
$\rightarrow Wcc$	1197	9.54
Zjj		
$Z \rightarrow \tau\tau$	110.2	0.25
$Z \rightarrow \nu\overline{\nu}$	2177	0.64
Zbb		
$Z \rightarrow \tau\tau$	5.31	0.53
$Z \rightarrow \nu\overline{\nu}$	190.8	20.7
Zcc		
$Z \rightarrow \tau\tau$	10.6	0.15
$Z \rightarrow \nu\overline{\nu}$	388.3	4.10
$t\bar{t}$	176.2	29.8
$Di - boson$	183.9	3.46
Total Physics Background	10150	113.1
Instrumental Background	2536	17.0
Total Background	12680	130.1
Observed Events	12490	140

TABLE I: Number of events after final selection.

	JES	Normalisation	Sideband fit	b-tagging	Sideband b-tagging	Trigger
Signal	5	10	3	7	-	6
Background	3	10	1	7	2	6

TABLE II: Table of systematic errors in %.

events. The expected signal is 1.4 events.

Higgs Mass (GeV)	$m_H = 105$	$m_H = 115$	$m_H = 125$	$m_H = 135$
ZH	1.04±0.16	0.84±0.13	0.60±0.09	0.35±0.05
Acceptance (%)	1.01	1.18	1.34	1.45
WH	0.69±0.10	0.56±0.08	0.39±0.06	0.23±0.03
Acceptance (%)	0.39	0.46	0.52	0.57
$Z + jets$	15.56	15.24	14.32	14.49
$W + jets$	24.87	24.82	24.23	24.73
$t\bar{t}$	13.01	15.33	17.07	18.97
Di-Boson	3.02	2.90	2.68	2.51
Instrumental Background	8.49	5.01	4.06	4.90
Expected Background	65.0±9.1	63.3±8.9	62.4±8.7	65.6±9.2
$N_{sig}/\sqrt{N_{bkgnd}}$	0.22	0.18	0.12	0.07
Observed Events	67	75	76	79
$ZH(H \rightarrow bb)$ Expected Limit (pb)	2.2	1.9	1.1	1.0
$ZH(H \rightarrow bb)$ Observed Limit (pb)	2.7	2.5	1.7	1.6
$WH(H \rightarrow bb)$ Expected Limit (pb)	5.5	5.0	2.7	2.5
$WH(H \rightarrow bb)$ Observed Limit (pb)	7.2	6.5	4.1	3.9

TABLE III: Number of events in the double b-tagged sample within a $\pm 2.0\sigma$ mass window. The cross section limits use the full invariant mass range.

As no significant excess is observed we proceed to set a limit on the Higgs production cross section. A modified frequentist approach with a Poisson log-likelihood ratio (LLR) statistic is used using the full invariant mass distribution as the discriminating variable [11] [12]. The impact of systematic uncertainties is incorporated through marginalisation of the Poisson probability distributions for signal and background via Gaussian distributions. All correlations in

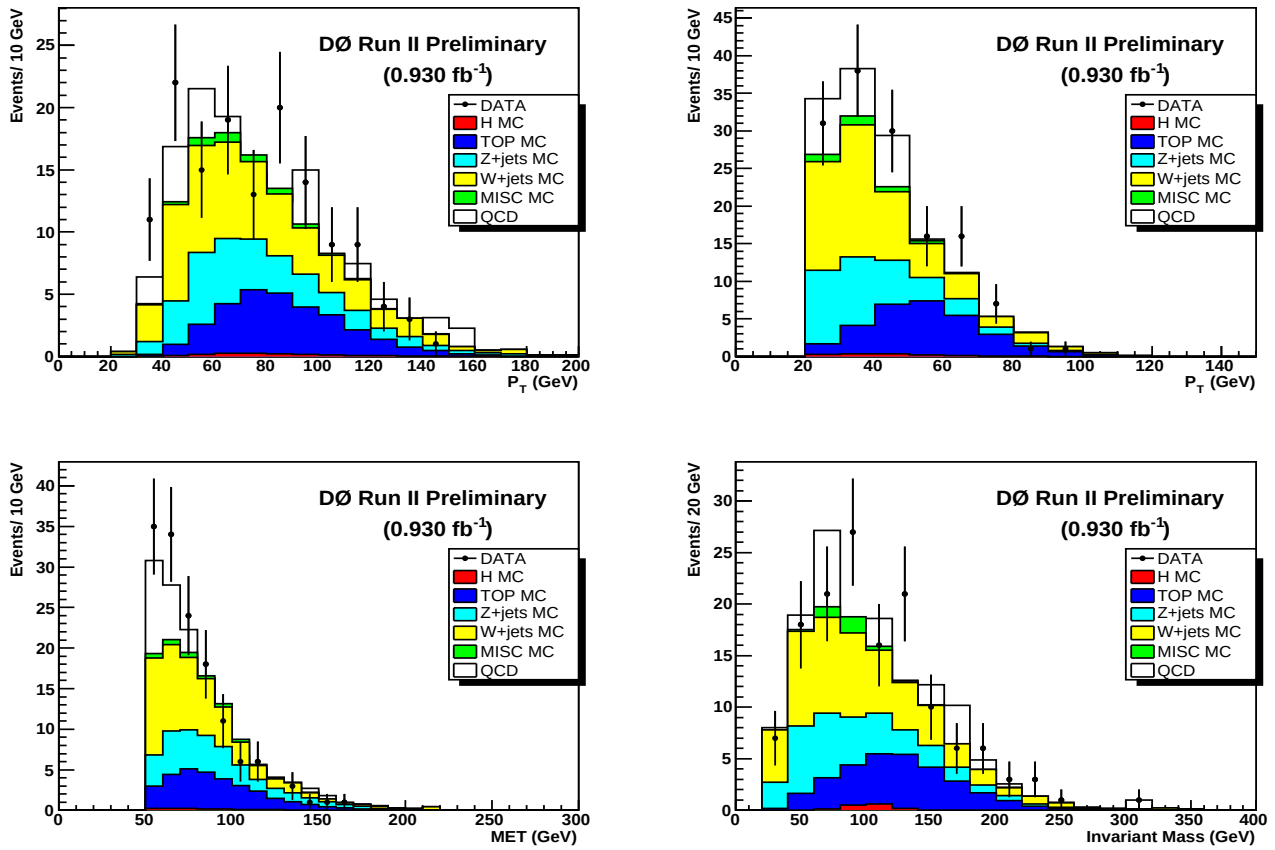

FIG. 4: Distributions of the p_T for the leading (upper left) and next-to-leading jet (upper right), $\not\!\!E_T$ and dijet mass after neural net b-tagging.

the systematic errors are maintained between signal and background. The expected distributions for background are evaluated by minimising a profile likelihood function, referencing the shape and rate of the distributions in the sideband regions. The observed and expected limits (pb) for the processes $\sigma(p\overline{p} \rightarrow ZH) \times (H \rightarrow b\overline{b})$ and $\sigma(p\overline{p} \rightarrow WH) \times (H \rightarrow b\overline{b})$ are given in Table III at 95% CL. The observed limits for a 115 GeV Higgs are 2.5 pb and 6.5 pb respectively, with expected limits of 1.9 pb and 5.0 pb. The expected and observed limits, for both the WH and ZH production, are displayed in Fig. 5, together with DØ's previously published result [1] and the Standard Model expectation.

Figure 6 shows the LLR distributions for the ZH and WH production in the $ZH \rightarrow \nu\overline{\nu}b\overline{b}$ channel. Included in these figures are the LLR values for the signal+background hypothesis (LLR_{s+b}), background-only hypothesis (LLR_b), and the observed data (LLR_{obs}). The shaded bands represent the 1 and 2 standard deviation (σ) departures for LLR_b. These distributions can be interpreted as follows:

- The separation between LLR_b and LLR_{s+b} provides a measure of the overall power of the search. This is the ability of the analysis to discriminate between the $s+b$ and $b-$only hypotheses.

- The width of the LLR_b distribution (shown here as 1 and 2 standard deviation (σ) bands) provides an estimate of how sensitive the analysis is to a signal-like fluctuation in data, taking account of the presence of systematic uncertainties. For example, when a 1-σ background fluctuation is large compared to the signal expectation, the analysis sensitivity is thereby limited.

- The value of LLR_{obs} relative to LLR_{s+b} and LLR_b indicates whether the data distribution appears to be more signal-like or background-like. As noted above, the significance of any departures of LLR_{obs} from LLR_b can be evaluated by the width of the LLR_b distribution.

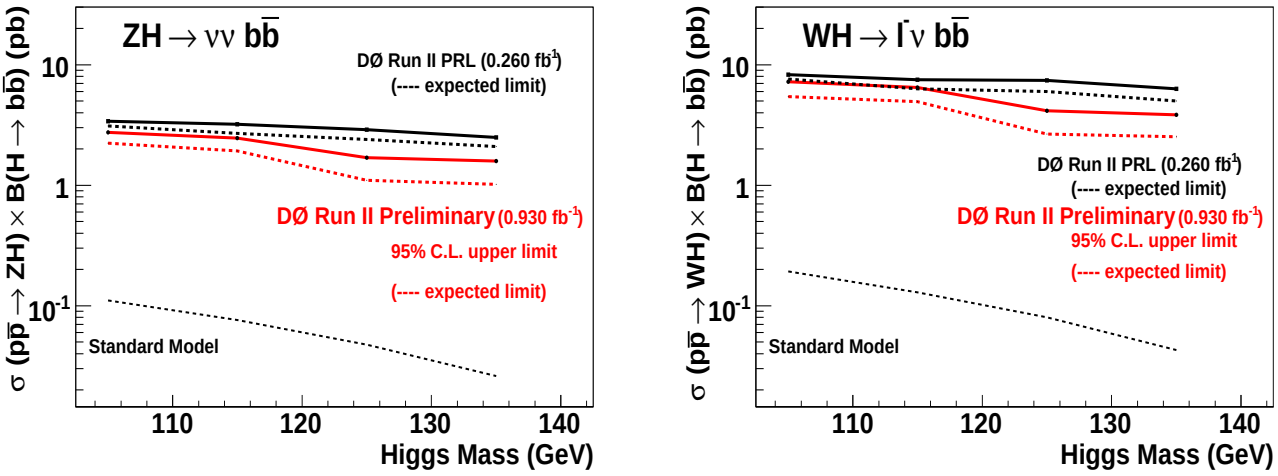

FIG. 5: Observed (solid) and expected (dashed) cross section limits at 95% CL on ZH production (left) and WH (right) production shown in red. Also shown is DØ's previously published result [1] (black) and the Standard Model cross section.

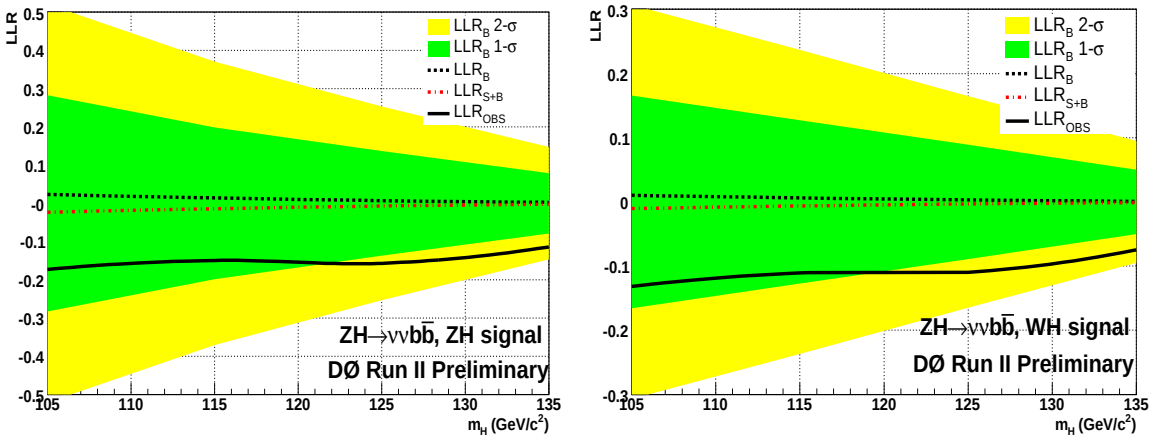

FIG. 6: LLR values for the signal+background hypothesis (LLR_{s+b}), background-only hypothesis (LLR_b), and the observed data (LLR_{obs}). The shaded bands represent the 1 and 2 standard deviation (σ) departures for LLR_b.

VIII. SUMMARY

We have performed a search for ZH associated production in the $\nu\bar{\nu}b\bar{b}$ channel using 0.930 fb^{-1} of data. As no significant excess is observed we set limits on the ZH production cross section of 2.7 pb – 1.6 pb at the 95% confidence level for Higgs boson masses from 105 – 135 GeV; the corresponding expected limits range from 2.2 pb to 1.0 pb.

Acknowledgments

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); PPARC (United Kingdom); MSMT (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); Research Corporation; Alexander von Humboldt Foundation; and the Marie

Curie Program.

---

[1] DØ Collaboration, Phys. Rev. Lett. **97**, 161803 (2006).
[2] LEP Collaborations, Phys. Lett. B **565**, 61 (2003).
[3] http://lepewwg.web.cern.ch/LEPEWWG/
[4] M. Carena *et al.*, "Report of the Tevatron Higgs Working Group', hep-ph/0010338.
[5] CDF and DØ Collaborations, "Results of the Tevatron Higgs Sensitivity Study", FERMILAB-PUB-03/320-E.
[6] DØ Collaboration, "The Upgraded DØ Detector", Nucl. Instrum. Methods Phys. Res. A **565**, 463 (2006).
[7] T. Andeen *et. al.*, FERMILAB-TM-2365-E (2006), in preparation.
[8] T. Sjostrand, L. Lonnblad, S. Mrenna and P. Skands,"PYTHIA 6.206", hep-ph/0108264.
[9] M. L. Mangano *et al.*, "ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions", hep-ph/0206293.
[10] T. Scanlon,"b-Tagging and the Search for Neutral Supersymmetric Higgs Bosons at DØ", FERMILAB-THESIS-2006-43.
[11] T. Junk, Nucl. Instrum. Methods A **434**, 435 (1999).
[12] W. Fisher, DØ Notes 4975 and 5309.
[13] Jets in the intercryostat region ($1.1 < |\eta| < 1.4$) have worse energy resolution and so are excluded from this analysis.
[14] For muons we require no objects with segments in the A and BC layer of the muon system [6] matched or not with a central track with $P_T > 8$ GeV and separated from any jet by $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\varphi^2} > 0.5$. Electrons are identified as having a calorimeter EM fraction of 0.9, the EM cluster must be matched to a track having $p_T > 5$ GeV and an 'electron likelihood' $>$ 0.2. If the isolation parameter, namely the ratio of the energy in a hollow cone with internal and external radii of R=0.2 and 0.4 respectively, is greater than 0.2 then the event is rejected.
[15] This sample has the same basic event selection as listed in Section II but in addition an isolated muon is required (as opposed to vetoing such events) along with an $\not\!\!E_T$ consistent with W production.